Supplemental Methods 1. Logistic regression model

Using a backwards elimination approach to arrive at a final model, multivariate logistic regression was developed to identify the variables that were associated with UMFA concentrations >1 nmol/L in the study population. Due to the limited DF, some preliminary analytic decisions were made to ensure there were enough DF to model several variables simultaneously and to investigate pairwise interactions. Continuous variables were treated as such whenever possible, in conjunction with the consideration of either transformations or higher order terms when the assumption of linearity with the log odds of UMFA >1 nmol/L did not seem reasonable. Categorical variables with ≥2 categories were collapsed when the prevalence estimates were similar and the context logical. Once a final model was developed for the entire study population (≥1 y, model 1), a subsequent model was fit limited to fasted persons ≥20 y (model 2) using the variables included in model 1 and adding daily alcohol volume, which was only publically available among adults ≥20 y.

Therefore, the initial logistic model 1included all variables listed below with a bivariate Wald χ^2 statistic *P*-value ≤0.10. The following categorical variables were considered: race-ethnicity (Hispanics, non-Hispanic black, non-Hispanic white), sex (male, female), dietary folic acid supplements (yes, no), smoking status (serum cotinine ≥10, <10), and fasting status (≥8 h, <8 h). Total folic acid intake (diet + supplements), RBC folate, and BSA were entered as continuous variables without any transformations or higher order terms; age was modeled with the addition of a quadratic term to account for the lack of linearity between age and the log odds of UMFA >1 nmol/L.

After these variables were entered into the initial model, we used the Satterthwaite-adjusted F to determine the statistical significance. The least significant variable in the model was dropped from the next model as long as there was no evidence of data-based confounding (>30% change in the parameter estimates of remaining variables). This strategy was iterated until all the variables in the final model were statistically significant (Satterthwaite-adjusted F *P*-value ≤0.05). Lastly, pairwise interactions between at least one categorical variable in the final model were explored. To facilitate the interpretation of the effect of age, the final models which included age with a quadratic term were reported using a categorized version of the continuous variable.